Probability 1

David Carlson

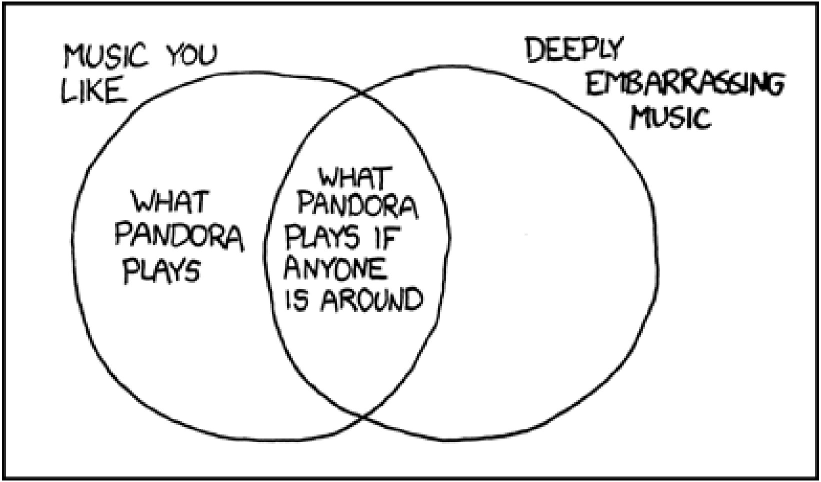
2021

Probability 1

David Carlson

2021

Probability 1


MUSIC YOU LIKE
DEEPLY EMBARRASSING MUSIC
WHAT PANDORA PLAYS
WHAT PANDORA PLAYS IF ANYONE IS AROUND

The big picture on probability

- In both formal modeling and statistics classes, you will constantly interact with concepts from basic probability theory

The big picture on probability

- In both formal modeling and statistics classes, you will constantly interact with concepts from basic probability theory
- In fact, many game theory and statistics texts will include small amounts of background information in passing

The big picture on probability

- In both formal modeling and statistics classes, you will constantly interact with concepts from basic probability theory
- In fact, many game theory and statistics texts will include small amounts of background information in passing
- This is often helpful, but can lead some students to get confused about what distinguishes statistical inference, the basic task in both statistics and game theoretic models under uncertainty, from probability theory

The big picture on probability

- In both formal modeling and statistics classes, you will constantly interact with concepts from basic probability theory
- In fact, many game theory and statistics texts will include small amounts of background information in passing
- This is often helpful, but can lead some students to get confused about what distinguishes statistical inference, the basic task in both statistics and game theoretic models under uncertainty, from probability theory
- So before you move forward into those areas of study, it is worth taking some time to understand probabilty by itself

Probability and inference

- The goal of inference is to take some observed data or known facts and backwards induct something about the world.

Probability and inference

- The goal of inference is to take some observed data or known facts and backwards induct something about the world.
  - For instance, we might want to survey a random subset of American citizens (our data) and estimate the true attitudes of the entire American electorate (the parameter)

Probability and inference

- The goal of inference is to take some observed data or known facts and backwards induct something about the world.
  - For instance, we might want to survey a random subset of American citizens (our data) and estimate the true attitudes of the entire American electorate (the parameter)
  - Alternatively, a game theoretic model may require actors to estimate the location of themedian voter given the sequence of prior election outcomes $x = (x_1, x_2, \ldots, x_n)$ and candidate positions $y = (y_1, y_2, \ldots, y_n)$.

- Probability theory is exactly the reverse

- Probability theory is exactly the reverse
  - Here we *know* the basic features of the data generating process (the parameters) and want to understand what the data is likely to look like

- Probability theory is exactly the reverse
  - Here we *know* the basic features of the data generating process (the parameters) and want to understand what the data is likely to look like
  - For instance, we might have a fair coin and we want to understand the likelihood of flipping 20 heads before th first tail shows up

- Probability theory is exactly the reverse
  - Here we *know* the basic features of the data generating process (the parameters) and want to understand what the data is likely to look like
  - For instance, we might have a fair coin and we want to understand the likelihood of flipping 20 heads before th first tail shows up
- Obviously, most of the things you are going to be doing in graduate school will be about inference. Nonetheless, you *really* needto have a grasp of probabilty theory first

What is probability theory?

- Probability is essentially thinking clearly about counting

What is probability theory?

- Probability is essentially thinking clearly about counting
- For the simplest problems, all you need to know is the number of ways that some set of outcomes X could happen versus the total number of ways things could have turned out

What is probability theory?

- Probability is essentially thinking clearly about counting
- For the simplest problems, all you need to know is the number of ways that some set of outcomes X could happen versus the total number of ways things could have turned out
- So to begin with, you just need to focus on getting a handle on the basic concepts of:

What is probability theory?

- Probability is essentially thinking clearly about counting
- For the simplest problems, all you need to know is the number of ways that some set of outcomes X could happen versus the total number of ways things could have turned out
- So to begin with, you just need to focus on getting a handle on the basic concepts of:
  - How to count events

What is probability theory?

- Probability is essentially thinking clearly about counting
- For the simplest problems, all you need to know is the number of ways that some set of outcomes X could happen versus the total number of ways things could have turned out
- So to begin with, you just need to focus on getting a handle on the basic concepts of:
  - How to count events
  - How to think about and handle sets

What is probability theory?

- Probability is essentially thinking clearly about counting
- For the simplest problems, all you need to know is the number of ways that some set of outcomes X could happen versus the total number of ways things could have turned out
- So to begin with, you just need to focus on getting a handle on the basic concepts of:
  - How to count events
  - How to think about and handle sets
  - How counting and sets relate to the concept of "probability"

What is probability theory?

- Probability is essentially thinking clearly about counting
- For the simplest problems, all you need to know is the number of ways that some set of outcomes X could happen versus the total number of ways things could have turned out
- So to begin with, you just need to focus on getting a handle on the basic concepts of:
  - How to count events
  - How to think about and handle sets
  - How counting and sets relate to the concept of “probability”
  - Conditional probability, independence, and Bayes’ law

Advanced Counting

A preliminary

- For $n \in \mathbb{N}$,

$$n! = \prod_{k=1}^{n} k$$

Advanced Counting

A preliminary

- For $n \in \mathbb{N}$,
$$n! = \prod_{k=1}^{n} k$$
- Example: $3! = 3 \cdot 2 \cdot 1 = 6$

Advanced Counting

A preliminary

- For $n \in \mathbb{N}$,

$$n! = \prod_{k=1}^{n} k$$

- Example: $3! = 3 \cdot 2 \cdot 1 = 6$

Advanced Counting

A preliminary

- For $n \in \mathbb{N}$,
$$n! = \prod_{k=1}^{n} k$$
- Example: $3! = 3 \cdot 2 \cdot 1 = 6$
- Example: $0! = 1$

Advanced Counting

A preliminary

- For $n \in \mathbb{N}$,
$$n! = \prod_{k=1}^{n} k$$
- Example: $3! = 3 \cdot 2 \cdot 1 = 6$
- Example: $0! = 1$

Advanced Counting

A preliminary

- For $n \in \mathbb{N}$,
$$n! = \prod_{k=1}^{n} k$$
- Example: $3! = 3 \cdot 2 \cdot 1 = 6$
- Example: $0! = 1$
- Example: for $x \geq y$,

Advanced Counting

A preliminary

- For $n \in \mathbb{N}$,
$$n! = \prod_{k=1}^{n} k$$
- Example: $3! = 3 \cdot 2 \cdot 1 = 6$
- Example: $0! = 1$
- Example: for $x \geq y$,

Advanced Counting

A preliminary

- For $n \in \mathbb{N}$,

$$n! = \prod_{k=1}^{n} k$$

- Example: $3! = 3 \cdot 2 \cdot 1 = 6$
- Example: $0! = 1$
- Example: for $x \geq y$,

$$\frac{x!}{(x-y)!} = \frac{x(x-1) \cdot \ldots \cdot (x-y+1) \cdot (x-y) \cdot (x-y-1) \cdot \ldots \cdot 1}{(x-y) \cdot (x-y-1) \cdot \ldots \cdot 1}$$

Advanced Counting

A preliminary

- For $n \in \mathbb{N}$,
$$n! = \prod_{k=1}^{n} k$$
- Example: $3! = 3 \cdot 2 \cdot 1 = 6$
- Example: $0! = 1$
- Example: for $x \geq y$,

$$\begin{aligned} \frac{x!}{(x-y)!} &= \frac{x(x-1) \cdot \ldots \cdot (x-y+1) \cdot (x-y) \cdot (x-y-1) \cdot \ldots \cdot 1}{(x-y) \cdot (x-y-1) \cdot \ldots \cdot 1} \\ &= x \cdot (x-1) \cdot \ldots \cdot (x-y+1) \end{aligned}$$

Fundamental Theorem of Counting

- If there are k characteristics, each with n_k alternatives, there are $\prod_{i=1}^{k} n_k$ possible outcomes
- We often need to count the number of ways to choose a subset from some set of possiblities. The number of outcomes depends on two characteristics of the process: does the *order* matter and is *replacement* allowed?

- If there are n objects and we select $k < n$ of them, how many different outcomes are possible?

- If there are n objects and we select $k < n$ of them, how many different outcomes are possible?
    1. Ordered, with replacement: n^k

- If there are n objects and we select $k < n$ of them, how many different outcomes are possible?
    1. Ordered, with replacement: n^k
    2. Ordered, without replacement: $\frac{n!}{(n-k)!}$

- If there are n objects and we select $k < n$ of them, how many different outcomes are possible?
  1. Ordered, with replacement: n^k
  2. Ordered, without replacement: $\frac{n!}{(n-k)!}$
  3. Unordered, with replacement: $\frac{(n+k-1)!}{(n-1)!k!} = \left(\begin{array}{c} n+k-1 \\ k \end{array} \right)$

- If there are n objects and we select $k < n$ of them, how many different outcomes are possible?
  1. Ordered, with replacement: n^k
  2. Ordered, without replacement: $\frac{n!}{(n-k)!}$
  3. Unordered, with replacement: $\frac{(n+k-1)!}{(n-1)!k!} = \binom{n+k-1}{k}$
  4. Unordered, without replacement: (n choose k): $\frac{n!}{(n-k)!k!} = \binom{n}{k}$

- Ordered events are sometimes referred to as permutations, while unordered events are combinations

- Ordered events are sometimes referred to as permutations, while unordered events are combinations
- In your introductory work, you will almost always be working with combinations

Thai Pizza Company menu

Select “Style: STIR FRIED (or) ON THE TOP”.

Thai Pizza Company menu

Select “Style: STIR FRIED (or) ON THE TOP”. Then, select:

COOKING

1) Pad Thai

2) Pad Kee Maow

3) Pad Num Prik Paow

4) Pad Num Prik Ong

5) Pad Num Prik Ong

6) Pad See Eew

7) Pad Spicy Sauce

8) Pad Sweet& Sour Sauce

9) Pad Green Curry

10) Pad Red Curry

11) Pad Gra Ree Curry

12) Pad Masaman Curry

13) Pad Panang Curry

14) Pad Peanut Sauce

Thai Pizza Company menu

Select “Style: STIR FRIED (or) ON THE TOP”. Then, select:

COOKING

1) Pad Thai
2) Pad Kee Maow
3) Pad Num Prik Paow
4) Pad Num Prik Ong
5) Pad Num Prik Ong
6) Pad See Eew
7) Pad Spicy Sauce
8) Pad Sweet& Sour Sauce
9) Pad Green Curry
10) Pad Red Curry
11) Pad Gra Ree Curry
12) Pad Masaman Curry
13) Pad Panang Curry
14) Pad Peanut Sauce

NOODLES

A) Rice Noodle (Small)
B) Rice Noodle (Large)
C) Vermicelli
D) Glass Noodle
E) Egg Noodle
F) Fresh Rice Noodle
G) Spaghetti
H) U-Dong Noodle
I) Soba Noodle
J) Mama Noodle
K) STEAM RICE

Thai Pizza Company menu

Select “Style: STIR FRIED (or) ON THE TOP”. Then, select:

COOKING

1) Pad Thai
2) Pad Kee Maow
3) Pad Num Prik Paow
4) Pad Num Prik Ong
5) Pad Num Prik Ong
6) Pad See Eew
7) Pad Spicy Sauce
8) Pad Sweet& Sour Sauce
9) Pad Green Curry
10) Pad Red Curry
11) Pad Gra Ree Curry
12) Pad Masaman Curry
13) Pad Panang Curry
14) Pad Peanut Sauce

NOODLES

A) Rice Noodle (Small)
B) Rice Noodle (Large)
C) Vermicelli
D) Glass Noodle
E) Egg Noodle
F) Fresh Rice Noodle
G) Spaghetti
H) U-Dong Noodle
I) Soba Noodle
J) Mama Noodle
K) STEAM RICE

MEAT

Chicken

Beef

Pork

Shrimp

Squid

Fish

Massel

Fish Ball

Shrimp Ball

Fish Cake

Tofu

Mushroom

Thai Pizza Company menu

Select “Style: STIR FRIED (or) ON THE TOP”. Then, select:

COOKING

1) Pad Thai
2) Pad Kee Maow
3) Pad Num Prik Paow
4) Pad Num Prik Ong
5) Pad Num Prik Ong
6) Pad See Eew
7) Pad Spicy Sauce
8) Pad Sweet& Sour Sauce
9) Pad Green Curry
10) Pad Red Curry
11) Pad Gra Ree Curry
12) Pad Masaman Curry
13) Pad Panang Curry
14) Pad Peanut Sauce

NOODLES

A) Rice Noodle (Small)
B) Rice Noodle (Large)
C) Vermicelli
D) Glass Noodle
E) Egg Noodle
F) Fresh Rice Noodle
G) Spaghetti
H) U-Dong Noodle
I) Soba Noodle
J) Mama Noodle
K) STEAM RICE

MEAT

Chicken

Beef

Pork

Shrimp

Squid

Fish

Massel

Fish Ball

Shrimp Ball

Fish Cake

Tofu

Mushroom

SPICE

*

**

***

****

*****

How long could I get lunch and never order the same thing twice?

How long could I get lunch and never order the same thing twice?

$$\prod_{k=1}^{K} n_k \quad = \quad 2 \cdot 14 \cdot 11 \cdot 12 \cdot 5 \tag{1}$$

How long could I get lunch and never order the same thing twice?

$$\prod_{k=1}^{K} n_k = 2 \cdot 14 \cdot 11 \cdot 12 \cdot 5 \tag{1}$$

$$= 18480 \quad \text{lunches} \tag{2}$$

How long could I get lunch and never order the same thing twice?

$$\prod_{k=1}^{K} n_k = 2 \cdot 14 \cdot 11 \cdot 12 \cdot 5 \quad (1)$$

$$= 18480 \quad \text{lunches} \quad (2)$$

$$\approx 51 \quad \text{years} \quad (3)$$

Sets

- **Set**: A set is any well defined collection of elements. If x is an element of S, $x \in S$.

Types of sets

1. Countably finite: a set with a finite number of elements, which can be mapped onto positive integers.

$$S = \{1, 2, 3, 4, 5, 6\}$$

2. Countably infinite: a set with an infinite number of elements, which can still be mapped onto positive integers.

$$S = \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$$

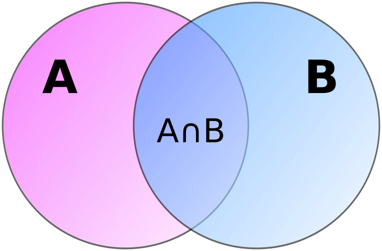
3. Uncountably infinite: a set with an infinite number of elements, which cannot be mapped onto positive integers.

$$S = \{x : x \in [0, 1]\}$$

4. Empty: a set with no elements.

$$S = \{\emptyset\}$$

Set operations



- **Union**: The union of two sets A and B, $A \cup B$, is the set containing all of the elements in A or B.
- **Intersection**: The intersection of sets A and B, $A \cap B$, is the set containing all of the elements in both A and B.
- **Complement**: If set A is a subset of S, then the complement of A, denoted A^C, is the set containing all of the elements in S that are not in A.

Properties of set operations:

1. Commutative: $A \cup B = B \cup A$, $A \cap B = B \cap A$
2. Associative: $A \cup (B \cup C) = (A \cup B) \cup C$, $A \cap (B \cap C) = (A \cap B) \cap C$
3. Distributive: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
4. de Morgan's laws: $(A \cup B)^C = A^C \cap B^C$, $(A \cap B)^C = A^C \cup B^C$

Disjointedness

- Sets are disjoint when they do not intersect, such that $A \cap B = \{\emptyset\}$. A collection of sets is pairwise disjoint if, for all $i \neq j$, $A_i \cap A_j = \{\emptyset\}$.
- A collection of sets form a partition of set S if they are pairwise disjoint and they cover set S, such that $\bigcup_{i=1}^{k} A_i = S$.

Let $A = 1, 5, 10$ and $B = 1, 2, \ldots, 10$

1. Is $A \subset B$, $B \subset A$, both, or neither?
2. What is $A \cup B$?
3. What is $A \cap B$?
4. Partition B into two sets, A and everything else. Call everything else C. What is C?
5. What is $A \cup C$?
6. What is $A \cap C$?

7. How many possible committees of 5 could be formed by 100 Senators?
8. If I gave people a survey with 5 questions (asked in a random order) with 4 response options for each question, how many combinations of responses could I get?
9. If I gave people a survey with 5 questions (asked in a random order) with 4 response options for each question, how many permutations of responses could I get?

Compute each of the followign

10.

$$\frac{12!}{7!}$$

11.

$$\frac{5!}{6!}$$

12.

$$\binom{12}{5}$$

Probability

- Probability is an expression of uncertainty

Probability

- Probability is an expression of uncertainty
- Modern probability theory is a way of estimating our uncertainty about some future events given specific assumed properties of the world

Probability

- Probability is an expression of uncertainty
- Modern probability theory is a way of estimating our uncertainty about some future events given specific assumed properties of the world
- This is a formalization of basic human intuition about how to handle risk

Sample Space

- A set or collection of all possible outcomes from some process

Sample Space

- A set or collection of all possible outcomes from some process
- Outcomes in the set can be discrete elements (countable) or points along a continuous interval (uncountable)

Sample Space

- A set or collection of all possible outcomes from some process
- Outcomes in the set can be discrete elements (countable) or points along a continuous interval (uncountable)
- Examples:

Sample Space

- A set or collection of all possible outcomes from some process
- Outcomes in the set can be discrete elements (countable) or points along a continuous interval (uncountable)
- Examples:
  1. **Discrete**: the numbers on a die, the number of possible wars that could occur each year, whether a vote cast is republican or democrat

Sample Space

- A set or collection of all possible outcomes from some process
- Outcomes in the set can be discrete elements (countable) or points along a continuous interval (uncountable)
- Examples:
  1. **Discrete**: the numbers on a die, the number of possible wars that could occur each year, whether a vote cast is republican or democrat
  2. **Continuous**: GNP, arms spending, age

Probability Distribution/Function

- A probability *function* on a sample space S is a mapping $\Pr(A)$ from events in S to the real numbers

Probability Distribution/Function

- A probability *function* on a sample space S is a mapping $\Pr(A)$ from events in S to the real numbers
- It is just like any other function

Probability Distribution/Function

- A probability *function* on a sample space S is a mapping $\Pr(A)$ from events in S to the real numbers
- It is just like any other function
- We have some event/sample space S we have a probability space (e.g., the probability of event x happening is some number in $[0, 1]$) and we have the function that translates x into the probability space that we denote $p(x)$ or $f(x)$

Example

- Let's say we are flipping two coins
- The **sample space** is $S = HH, HT, TH, TT$
- We are interested in the number of heads

Example

- Let's say we are flipping two coins
- The **sample space** is $S = HH, HT, TH, TT$
- We are interested in the number of heads

Outcome	X=x
HH	2
HT	1
TH	1
TT	0

Example

- Let's say we are flipping two coins
- The **sample space** is $S = HH, HT, TH, TT$
- We are interested in the number of heads

Outcome	X=x
HH	2
HT	1
TH	1
TT	0

x	p(x)
TT $\rightarrow$ 0	.25
HT, TH $\rightarrow$ 1	.50
HH $\rightarrow$ 2	.25
Sum	1.00

Exercise: Rolling two fair dice

1. Write out the sample space
2. Write out the empirical probability function

Axioms of Probability

- Probability functions will satisfy the following three axioms (due to Kolmogorov).
- Define the number $\Pr(A)$ correponding to each event A in the sample space S such that:

  1. Axiom: For any event A, $\Pr(A) \geq 0$.
  2. Axiom: $\Pr(S) = 1$
  3. Axiom: For any sequence of disjoint events $A_1, A_2, \ldots$ (of which there may be infinitely many),

$$\Pr\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} \Pr(A_i)$$

Basic Theorems of Probability

- Using these three axioms, we can define all of the common theorems of probability.
  1. $\Pr(\emptyset) = 0$
  2. $\Pr(A^C) = 1 - \Pr(A)$
  3. For any event A, $0 \leq \Pr(A) \leq 1$.

4. If $A \subset B$, then $\Pr(A) \leq \Pr(B)$.
5. For any two events A and B,
   $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$
6. For any sequence of n events (which need not be disjoint)

$$A_1, A_2, \ldots, A_n$$

,

$$\Pr\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \Pr(A_i)$$

Example 1

Let's assume we have an evenly-balanced, six-sided die. Then,

1. Sample space $S = \{1, 2, 3, 4, 5, 6\}$
2. $\Pr(1) = \cdots = \Pr(6) = 1/6$
3. $\Pr(\emptyset) = \Pr(7) = 0$
4. $\Pr(\{1, 3, 5\}) = 1/6 + 1/6 + 1/6 = 1/2$

5. $\Pr\left(\overline{\{1,2\}}\right) = \Pr\left(\{3,4,5,6\}\right) = 2/3$
6. Let $B = S$ and $A = \{1,2,3,4,5\} \subset B$. Then $\Pr(A) = 5/6 < \Pr(B) = 1$.
7. Let $A = \{1,2,3\}$ and $B = \{2,4,6\}$. Then $A \cup B = \{1,2,3,4,6\}$, $A \cap B = \{2\}$, and

$$\begin{aligned}
\Pr(A \cup B) &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \\
&= 3/6 + 3/6 - 1/6 \\
&= 5/6
\end{aligned}$$

Example 2

Let trial be equally likely selection of one of X, Y, Z for exit poll.

Example 2

Let trial be equally likely selection of one of X, Y, Z for exit poll.

1. What's the sample space?

Example 2

Let trial be equally likely selection of one of X, Y, Z for exit poll.

1. What's the sample space? $\{X, Y, Z\}$

Example 2

Let trial be equally likely selection of one of X, Y, Z for exit poll.

1. What's the sample space? $\{X, Y, Z\}$
2. For event $A =$ "select X", calculate $p(A)$.

Example 2

Let trial be equally likely selection of one of X, Y, Z for exit poll.

1. What's the sample space? $\{X, Y, Z\}$
2. For event $A =$ "select X", calculate $p(A)$. $p(X) = \frac{1}{3}$
3. Calculate $p(A^C)$.

Example 2

Let trial be equally likely selection of one of X, Y, Z for exit poll.

1. What's the sample space? $\{X, Y, Z\}$
2. For event $A =$ "select X", calculate $p(A)$. $p(X) = \frac{1}{3}$
3. Calculate $p(A^C)$. $p(Y \cup Z) = \frac{2}{3}$

Example 2

Let trial be equally likely selection of one of X, Y, Z for exit poll.

1. What's the sample space? $\{X, Y, Z\}$
2. For event $A =$ "select X", calculate $p(A)$. $p(X) = \frac{1}{3}$
3. Calculate $p(A^C)$. $p(Y \cup Z) = \frac{2}{3}$
4. Calculate $p(A \cup A^C)$.

Example 2

Let trial be equally likely selection of one of X, Y, Z for exit poll.

1. What's the sample space? $\{X, Y, Z\}$
2. For event $A =$ "select X", calculate $p(A)$. $p(X) = \frac{1}{3}$
3. Calculate $p(A^C)$. $p(Y \cup Z) = \frac{2}{3}$
4. Calculate $p(A \cup A^C)$. $p(X \cup (Y \cup Z)) = 1$

Conditional Probability and Bayes' Law

- Let A, B be two events with $p(A) > 0$,

Conditional Probability and Bayes' Law

- Let A, B be two events with $p(A) > 0$,
  - the *conditional probability* of B given A is

$$p(B|A) = \frac{p(B \cap A)}{p(A)}$$

Conditional Probability and Bayes' Law

- Let A, B be two events with $p(A) > 0$,
  - the *conditional probability* of B given A is

$$p(B|A) = \frac{p(B \cap A)}{p(A)}$$

  - Equivalently,

$$p(B \cap A) = p(A)p(B|A)$$

Conditional Probability and Bayes' Law

- Let A, B be two events with $p(A) > 0$,
  - the *conditional probability* of B given A is

$$p(B|A) = \frac{p(B \cap A)}{p(A)}$$

  - Equivalently,

$$p(B \cap A) = p(A)p(B|A)$$

  - Relatedly,

$$p(A \cap B \cap C) = p(A)p(B|A)p(C|A \cap B)$$

Conditional Probability and Bayes' Law

- Let A, B be two events with $p(A) > 0$,
  - the *conditional probability* of B given A is
$$p(B|A) = \frac{p(B \cap A)}{p(A)}$$
  - Equivalently,
$$p(B \cap A) = p(A)p(B|A)$$
  - Relatedly,
$$p(A \cap B \cap C) = p(A)p(B|A)p(C|A \cap B)$$
  - Conditioning information can be subtlely important

Example: Older Child Paradox

- Consider picking a two-child family at random, with equiprobable $S = \{FF, MM, FM, MF\}$.

Example: Older Child Paradox

- Consider picking a two-child family at random, with equiprobable $S = \{FF, MM, FM, MF\}$.
- Let $B = FF$, so $p(B) = .25$.

Example: Older Child Paradox

- Consider picking a two-child family at random, with equiprobable $S = \{FF, MM, FM, MF\}$.
- Let $B = FF$, so $p(B) = .25$.
- Let A be "at least one girl". Calculate $p(B|A)$.

Example: Older Child Paradox

- Consider picking a two-child family at random, with equiprobable $S = \{FF, MM, FM, MF\}$.
- Let $B = FF$, so $p(B) = .25$.
- Let A be "at least one girl". Calculate $p(B|A)$.

$$p(B|A) = \frac{p(B \cap A)}{p(A)} = \frac{1/4}{3/4} = \frac{1}{3}$$

- Let C be "older child is girl". Calculate $p(B|C)$.

Example: Older Child Paradox

- Consider picking a two-child family at random, with equiprobable $S = \{FF, MM, FM, MF\}$.
- Let $B = FF$, so $p(B) = .25$.
- Let A be "at least one girl". Calculate $p(B|A)$.

$$p(B|A) \quad = \quad \frac{p(B \cap A)}{p(A)} = \frac{1/4}{3/4} = \frac{1}{3}$$

- Let C be "older child is girl". Calculate $p(B|C)$.

$$p(B|C) \quad = \quad \frac{p(B \cap C)}{p(C)} = \frac{1/4}{2/4} = \frac{1}{2}$$

Example: The Birthday Problem

What's the probability that two classmates have the same birthday?

- 2 students: 1 has a birthday.

Example: The Birthday Problem

What's the probability that two classmates have the same birthday?

- 2 students: 1 has a birthday. $p(match) = \frac{1}{365} \approx .003$

Example: The Birthday Problem

What's the probability that two classmates have the same birthday?

- 2 students: 1 has a birthday. $p(match) = \frac{1}{365} \approx .003$

- 3 students.

- 3 students.

$$p(match_{12}) = \frac{1}{365}$$

- 3 students.

$$p(match_{12}) = \frac{1}{365}$$

$$p(match_{23}) = \frac{1}{365}$$

$$p(match_{13}) = \frac{1}{365}$$

  - Let $A = match_{12}$, $B = match_{13}$, $C = match_{23}$.

- 3 students.

$$p(match_{12}) = \frac{1}{365}$$

$$p(match_{23}) = \frac{1}{365}$$

$$p(match_{13}) = \frac{1}{365}$$

  - Let $A = match_{12}$, $B = match_{13}$, $C = match_{23}$.

$$\begin{aligned} p(A \cup B \cup C) &= p(A) + p(B) + p(C) \\ &\quad - p(A \cap B) - p(A \cap C) - p(B \cap C) + p(A \cap B \cap C) \end{aligned}$$

- 3 students.

$$p(match_{12}) = \frac{1}{365}$$

$$p(match_{23}) = \frac{1}{365}$$

$$p(match_{13}) = \frac{1}{365}$$

  - Let $A = match_{12}$, $B = match_{13}$, $C = match_{23}$.

$$\begin{aligned} p(A \cup B \cup C) &= p(A) + p(B) + p(C) \\ &\quad - p(A \cap B) - p(A \cap C) - p(B \cap C) + p(A \cap B \cap C) \\ &= \frac{1}{365} + \frac{1}{365} + \frac{1}{365} - (\frac{1}{365} \cdot \frac{1}{365}) \\ &\quad - (\frac{1}{365} \cdot \frac{1}{365}) - (\frac{1}{365} \cdot \frac{1}{365}) + (\frac{1}{365} \cdot \frac{1}{365} \cdot 1) \end{aligned}$$

- 3 students.

$$p(match_{12}) = \frac{1}{365}$$

$$p(match_{23}) = \frac{1}{365}$$

$$p(match_{13}) = \frac{1}{365}$$

  - Let $A = match_{12}$, $B = match_{13}$, $C = match_{23}$.

$$\begin{aligned}
p(A \cup B \cup C) &= p(A) + p(B) + p(C) \\
&\quad - p(A \cap B) - p(A \cap C) - p(B \cap C) + p(A \cap B \cap C) \\
&= \frac{1}{365} + \frac{1}{365} + \frac{1}{365} - (\frac{1}{365} \cdot \frac{1}{365}) \\
&\quad -(\frac{1}{365} \cdot \frac{1}{365}) - (\frac{1}{365} \cdot \frac{1}{365}) + (\frac{1}{365} \cdot \frac{1}{365} \cdot 1) \\
&\approx .008 \qquad (4)
\end{aligned}$$

An easier way: Use complements.

$$p(match) = 1 - p(nomatch)$$

An easier way: Use complements.

$$\begin{aligned} p(\mathit{match}) &= 1 - p(\mathit{nomatch}) \\ p(\mathit{match}_{123}) &= 1 - p(\mathit{nomatch}_{123}) \end{aligned}$$

An easier way: Use complements.

$$\begin{aligned}
p(match) &= 1 - p(nomatch) \\
p(match_{123}) &= 1 - p(nomatch_{123}) \\
&= 1 - p(nomatch_{12}) \cdot p(nomatch_{13} | nomatch_{12}) \\
&\quad \cdot p(nomatch_{23} | nomatch_{12} \cap nomatch_{13})
\end{aligned}$$

An easier way: Use complements.

$$
\begin{aligned}
p(match) &= 1 - p(nomatch) \\
p(match_{123}) &= 1 - p(nomatch_{123}) \\
&= 1 - p(nomatch_{12}) \cdot p(nomatch_{13} | nomatch_{12}) \\
&\quad \cdot p(nomatch_{23} | nomatch_{12} \cap nomatch_{13}) \\
&= 1 - \frac{364}{365} \cdot \frac{364}{365} \cdot \frac{363}{364}
\end{aligned}
$$

An easier way: Use complements.

$$
\begin{aligned}
p(match) &= 1 - p(nomatch) \\
p(match_{123}) &= 1 - p(nomatch_{123}) \\
&= 1 - p(nomatch_{12}) \cdot p(nomatch_{13} | nomatch_{12}) \\
&\quad \cdot p(nomatch_{23} | nomatch_{12} \cap nomatch_{13}) \\
&= 1 - \frac{364}{365} \cdot \frac{364}{365} \cdot \frac{363}{364} \\
&\approx .008
\end{aligned}
$$

Example:

- A six-sided die is rolled.
- What is the probability of a 1, given the outcome is an odd number?
- Let $A = \{1\}$, $B = \{1, 3, 5\}$, and $A \cap B = \{1\}$.

Example:

- A six-sided die is rolled.
- What is the probability of a 1, given the outcome is an odd number?
- Let $A = \{1\}$, $B = \{1, 3, 5\}$, and $A \cap B = \{1\}$.
- Then, $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{1/6}{1/2} = 1/3$.

Multiplicative Law of Probability

- The probability of the intersection of two events A and B is

$$\Pr(A \cap B) = \Pr(A)\Pr(B|A) = \Pr(B)\Pr(A|B)$$

Multiplicative Law of Probability

- The probability of the intersection of two events A and B is

$$\Pr(A \cap B) = \Pr(A)\Pr(B|A) = \Pr(B)\Pr(A|B)$$

- Follows directly from the definition of conditional probability.

Law of Total Probability

- Let S be the sample space of some experiment and let the disjoint k events $B_1, \dots, B_k$ partition S.
- If A is some other event in S, then the events $AB_1, AB_2, \dots, AB_k$ will form a partition of A and we can write A as

$$A = (AB_1) \cup \cdots \cup (AB_k)$$

- Since the k events are disjoint,

$$\begin{aligned} \Pr(A) &= \sum_{i=1}^{k} \Pr(A, B_i) \\ &= \sum_{i=1}^{k} \Pr(B_i) \Pr(A|B_i) \end{aligned}$$

- Sometimes it is easier to calculate the conditional probabilities and sum them than it is to calculate $\Pr(A)$ directly.

Bayes' Law

- From conditional probability,

$$p(B \cap A) = p(A)p(B|A)$$

Bayes' Law

- From conditional probability,

$$p(B \cap A) = p(A)p(B|A)$$

- But also,

$$p(A \cap B) = p(B)p(A|B)$$

Bayes' Law

- From conditional probability,

$$p(B \cap A) = p(A)p(B|A)$$

- But also,

$$p(A \cap B) = p(B)p(A|B)$$

- Since we know $p(B \cap A) = p(A \cap B)$,

Bayes' Law

- From conditional probability,

$$p(B \cap A) = p(A)p(B|A)$$

- But also,

$$p(A \cap B) = p(B)p(A|B)$$

- Since we know $p(B \cap A) = p(A \cap B)$,

$$p(A)p(B|A) \quad = \quad p(B)p(A|B)$$

Bayes' Law

- From conditional probability,

$$p(B \cap A) = p(A)p(B|A)$$

- But also,

$$p(A \cap B) = p(B)p(A|B)$$

- Since we know $p(B \cap A) = p(A \cap B)$,

$$\begin{aligned} p(A)p(B|A) &= p(B)p(A|B) \\ p(B|A) &= \frac{p(B)p(A|B)}{p(A)} \end{aligned}$$

Then, from the Law of Total Probability,

$$p(B|A) = \frac{p(B)p(A|B)}{p(A)}$$

Then, from the Law of Total Probability,

$$
\begin{aligned}
p(B|A) &= \frac{p(B)p(A|B)}{p(A)} \\
&= \frac{p(B)p(A|B)}{p(A \cap B) + p(A \cap B^C)}
\end{aligned}
$$

Then, from the Law of Total Probability,

$$
\begin{aligned}
p(B|A) &= \frac{p(B)p(A|B)}{p(A)} \\
&= \frac{p(B)p(A|B)}{p(A \cap B) + p(A \cap B^C)} \\
&= \frac{p(B)p(A|B)}{p(B)p(A|B) + p(B^C)p(A|B^C)}
\end{aligned}
$$

More generally, where the B_j form a partition,

$$p(B_j|A) = \frac{p(B_j)p(A|B_j)}{\sum_{i=1}^{k} p(B_i)p(A|B_i)}$$

More generally, where the B_j form a partition,

$$p(B_j|A) \quad = \quad \frac{p(B_j)p(A|B_j)}{\sum_{i=1}^{k} p(B_i)p(A|B_i)}$$

These are *Bayes' Law* or *Bayes' Theorem* or *Bayes' Rule*.

Thinking about Bayes' Rule

- Assume that events $B_1, \ldots, B_k$ form a partition of the space S.
- Then

$$\Pr(B_j|A) = \frac{\Pr(A, B_j)}{\Pr(A)} = \frac{\Pr(B_j)\Pr(A|B_j)}{\sum_{i=1}^{k} \Pr(B_i)\Pr(A|B_i)}$$

- If there are only two states of B, then this is just

$$\Pr(B_1|A) = \frac{\Pr(B_1)\Pr(A|B_1)}{\Pr(B_1)\Pr(A|B_1) + \Pr(B_2)\Pr(A|B_2)}$$

- If this was a continuouse distribution we could write this as:

$$\Pr(B_j|A) = \frac{\Pr(A, B_j)}{\Pr(A)} = \frac{\Pr(B)\Pr(A|B)}{\int_{\infty}^{\infty} \Pr(A, B)\Pr(B)}$$

- Note that the denominator has an indefinite integral, meaning that there is an unknown integration constant to consider.

In Bayesian modeling and data analysis,

$$p(\theta|y) \quad \propto \quad p(\theta)p(y|\theta)$$

In Bayesian modeling and data analysis,

$$\begin{aligned} p(\theta|y) &\propto p(\theta)p(y|\theta) \\ \text{posterior} &\propto \text{prior} \cdot \text{likelihood} \end{aligned}$$

Example: Rare conditions and "accurate" tests

- A test for cancer correctly detects it 90% of the time, but incorrectly identifies a person as having cancer 10% of the time. If 10% of all people have cancer at any given time, what is the probability that a person who tests positive actually has cancer?

$$p(B|A) = \frac{p(B)p(A|B)}{p(B)p(A|B) + p(B^C)p(A|B^C)}$$

Example: Rare conditions and "accurate" tests

- A test for cancer correctly detects it 90% of the time, but incorrectly identifies a person as having cancer 10% of the time. If 10% of all people have cancer at any given time, what is the probability that a person who tests positive actually has cancer?

$$
\begin{aligned}
p(B|A) &= \frac{p(B)p(A|B)}{p(B)p(A|B) + p(B^C)p(A|B^C)} \\
p(Y|+) &= \frac{p(Y)p(+|Y)}{p(Y)p(+|Y) + p(Y^C)p(+|Y^C)}
\end{aligned}
$$

Example: Rare conditions and “accurate” tests

- A test for cancer correctly detects it 90% of the time, but incorrectly identifies a person as having cancer 10% of the time. If 10% of all people have cancer at any given time, what is the probability that a person who tests positive actually has cancer?

$$
\begin{aligned}
p(B|A) &= \frac{p(B)p(A|B)}{p(B)p(A|B) + p(B^C)p(A|B^C)} \\
p(Y|+) &= \frac{p(Y)p(+|Y)}{p(Y)p(+|Y) + p(Y^C)p(+|Y^C)} \\
&= \frac{.1 \cdot p(+|Y)}{.1 \cdot p(+|Y) + .9 \cdot p(+|Y^C)}
\end{aligned}
$$

Example: Rare conditions and “accurate” tests

- A test for cancer correctly detects it 90% of the time, but incorrectly identifies a person as having cancer 10% of the time. If 10% of all people have cancer at any given time, what is the probability that a person who tests positive actually has cancer?

$$
\begin{aligned}
p(B|A) &= \frac{p(B)p(A|B)}{p(B)p(A|B) + p(B^C)p(A|B^C)} \\
p(Y|+) &= \frac{p(Y)p(+|Y)}{p(Y)p(+|Y) + p(Y^C)p(+|Y^C)} \\
&= \frac{.1 \cdot p(+|Y)}{.1 \cdot p(+|Y) + .9 \cdot p(+|Y^C)} \\
&= \frac{.1 \cdot .9}{.1 \cdot .9 + .9 \cdot .1}
\end{aligned}
$$

Example: Rare conditions and "accurate" tests

- A test for cancer correctly detects it 90% of the time, but incorrectly identifies a person as having cancer 10% of the time. If 10% of all people have cancer at any given time, what is the probability that a person who tests positive actually has cancer?

$$
\begin{aligned}
p(B|A) &= \frac{p(B)p(A|B)}{p(B)p(A|B) + p(B^C)p(A|B^C)} \\
p(Y|+) &= \frac{p(Y)p(+|Y)}{p(Y)p(+|Y) + p(Y^C)p(+|Y^C)} \\
&= \frac{.1 \cdot p(+|Y)}{.1 \cdot p(+|Y) + .9 \cdot p(+|Y^C)} \\
&= \frac{.1 \cdot .9}{.1 \cdot .9 + .9 \cdot .1} \\
&= \frac{.09}{.09 + .09}
\end{aligned}
$$

Example: Rare conditions and "accurate" tests

- A test for cancer correctly detects it 90% of the time, but incorrectly identifies a person as having cancer 10% of the time. If 10% of all people have cancer at any given time, what is the probability that a person who tests positive actually has cancer?

$$
\begin{aligned}
p(B|A) &= \frac{p(B)p(A|B)}{p(B)p(A|B) + p(B^C)p(A|B^C)} \\
p(Y|+) &= \frac{p(Y)p(+|Y)}{p(Y)p(+|Y) + p(Y^C)p(+|Y^C)} \\
&= \frac{.1 \cdot p(+|Y)}{.1 \cdot p(+|Y) + .9 \cdot p(+|Y^C)} \\
&= \frac{.1 \cdot .9}{.1 \cdot .9 + .9 \cdot .1} \\
&= \frac{.09}{.09 + .09} \\
&= .5
\end{aligned}
$$

Class Exercise

- In Boston, 30% of the people are conservatives, 50% are liberals, and 20% are independents.
- In the last election, 65% of conservatives, 82% of liberals, and 50% of independents voted.
- If a person in Boston is selected at random and we learn that s/he did not vote last election, what is the probability s/he is a liberal?

1. If A, B, C, and D aare mutually exclusive and collectively exhaustive, what is the joint probability of A, B, C, AND D?
2. If A, B, C, and D aare mutually exclussive and collectively exhaustive, what is the joing probability of A, B, C, OR D?
3. Solve what is known as the Monte Hall problem. There are three doors. Behind two o these are goats, while behind the third is a new car. You choose one door. Monte Hall opens one of the other two doors, revealing a goat, and asks if you'd like to stick with the door you have, or switch to the other door he did not open. You get whatever is behind the door you chooose. Should you switch doors? Why or why not?

In a certain city, 30% of the citizens are conservatives, 30% are liberals, and 40% are independents. In a recent election, 50% of conservatives voted, 40% of liberals voted, and 30% of independents voted.

4. What is the probability that a person voted?
5. If the person voted, what is the probability that the voter is conservative?
6. Liberal?

7. In rolling two dice labeled X and Y, what is the probability that the sum of the up faces is four, given that either X or Y shows a three?

Use this joint probability distribution:

		X		
		0	1	2
	0	0.10	0.10	0.01
Y	1	0.02	0.10	0.20
	2	0.30	0.10	0.07

8. $p(X < 2)$
9. $p(X < 2|Y < 2)$
10. $\Pr(Y = 2|x \leq 1)$
11. $p(X = 1|Y = 1)$
12. $p(Y > 0|X > 0)$

13.

Assume that 2% of the population of the United States are members of some extremist militia group, ($p(M) = 0.02$). However, members may be unwilling to admit their mempership on a survey.

We develope a survey question that is 95% accurate on positive classification $p(C|M) = 0.95$ and 97% accurate on negative classification, $P(C^C|M^C) = 0.97$.

Using Bayes' Law, derive the probability that someone positively classified by the survey as being a militia member really is a militia member.

Independence

- If the occurrence or nonoccurrence of either events A and B have no effect on the occurrence or nonoccurrence of the other, then A and B are **independent**.
- If A and B are independent, then
  1. $\Pr(A|B) = \Pr(A)$
  2. $\Pr(B|A) = \Pr(B)$
  3. $\Pr(A \cap B) = \Pr(A)\Pr(B)$

Pairwise independence

- A set of more than two events $A_1, A_2, \ldots, A_k$ is **pairwise independent** if $\Pr(A_i \cap A_j) = \Pr(A_i)\Pr(A_j)$, $\forall i \neq j$.
- Note that this does *not* necessarily imply that $\Pr(\bigcap_{i=1}^{k} A_i) = \prod_{i=1}^{K} \Pr(A_i)$.

- Consider two flips of a fair coin.

- Consider two flips of a fair coin. $\{HH, HT, TH, TT\}$. Let
    1. $A_1 = H\square$
    2. $A_2 = \square H$
    3. $A_3 =$ exactly one H
- These are pairwise independent

- Consider two flips of a fair coin. $\{HH, HT, TH, TT\}$. Let
  1. $A_1 = H\square$
  2. $A_2 = \square H$
  3. $A_3 =$ exactly one H
- These are pairwise independent, but not independent as a group:

- Consider two flips of a fair coin. $\{HH, HT, TH, TT\}$. Let
  1. $A_1 = H\square$
  2. $A_2 = \square H$
  3. $A_3 =$ exactly one H
- These are pairwise independent, but not independent as a group:

$$p(A_1 \cap A_2 \cap A_3) \quad = \quad 0$$

- Consider two flips of a fair coin. $\{HH, HT, TH, TT\}$. Let
  1. $A_1 = H\square$
  2. $A_2 = \square H$
  3. $A_3 =$ exactly one H
- These are pairwise independent, but not independent as a group:

$$\begin{aligned} p(A_1 \cap A_2 \cap A_3) &= 0 \\ p(A_1)p(A_2)p(A_3) &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8} \end{aligned}$$

Conditional independence

- If the occurrence of A or B conveys no information about the occurrence of the other, once you know the occurrence of a third event C, then A and B are **conditionally independent** (conditional on C):
  1. $\Pr(A|B \cap C) = \Pr(A|C)$
  2. $\Pr(B|A \cap C) = \Pr(B|C)$
  3. $\Pr(A \cap B|C) = \Pr(A|C)\Pr(B|C)$
- Conditional independence is one of the fundamental assumptions deployed for most statistical estimation techniques. It is a *very* strong assumption.

If A and B are independent events, are the followign true or false?

1. $$\Pr(A \cap B) = \Pr(A)\Pr(B)$$
2. $$\Pr(A|B) = \Pr(A) + \Pr(A)\Pr(B)$$
3. $$\Pr(B|A) = \Pr(B)$$
4. Let $P(A) = 0.3$ and $P(A \cup B) = .5$. Find $P(B)$, assuming both events are independent?
5. What problems do you run into when they are not independent?